Prediction of Heart Disease

**Algorithms Used:**

* Logistic Regression
* K-Nearest Neighbors
* Random Forest

**Colab Notebook:** <https://colab.research.google.com/drive/1F3mUHkH1mRgErhpsGSTqRUJ2DofYOA9C?usp=sharing>

**Source:**

<https://www.youtube.com/watch?v=zoYxGeP5KW4>

**Dataset:**

<https://github.com/MuhammadAsifff/AIMachineLearningDataScienceProjectNA/blob/master/dataset_heart.csv>

**Rows, columns:** (303, 14)

**About:**

Understanding about heart health is usually a combination several factors such as age, gender, heart rate and several other factors that may be useful to detect and counter any future effects to the heart.

The analysis includes exploring data patterns, correlations between features like age, heart rate, and chest pain type, and building machine learning models for prediction. The models are evaluated and tuned to improve their accuracy. Evaluation metrics like accuracy, precision, recall, and F1-score assess model performance. Feature importance is also analyzed to understand the impact of different factors on heart disease prediction.

Parameters:

* age: Age of the patient.
* sex: Gender of the patient (1 = male; 0 = female).
* chest pain type (cp): Type of chest pain (4 values: 0, 1, 2, 3). Value 0: typical angina Value 1: atypical angina Value 2: non-anginal pain Value 3: asymptomatic
* resting blood pressure (trestbps/blood\_pressure): Resting blood pressure in mm Hg.
* serum cholesterol (chol/cholesterol): Serum cholesterol in mg/dl.
* fasting blood sugar (fbs/blood\_sugar): Fasting blood sugar > 120 mg/dl (1 = true; 0 = false).
* resting electrocardiographic results (restecg): Resting electrocardiographic results (values 0, 1, 2).
* maximum heart rate achieved (thalach): Maximum heart rate achieved.
* exercise-induced angina (exang): Exercise-induced angina (1 = yes; 0 = no).
* oldpeak: ST depression induced by exercise
* slope: The slope of the peak exercise ST segment.
* number of major vessels (ca/vessels): Number of major vessels (0-3) colored by fluoroscopy.
* thal: 3 = normal; 6 = fixed defect; 7 = reversible defect (thallium heart scan or stress test).
* target: Target variable (0 = no heart disease; 1 = heart disease).

**Algorithms Model and Accuracy**

|  |  |
| --- | --- |
| **Models** | **Scores** |
| Logistic Regression | 0.7540 |
| K-Nearest Neighbors | 0.7049 |
| Random Forest | 0.7049 |

Best: **Logistic Regression**